

DETAILED PROOFS (CLASS NOTES) OF NECESSARY AND SUFFICIENT CONDITIONS FOR
INDEPENDENCE AND χ^2 -DISTRIBUTION PROPERTIES OF QUADRATIC FORMS
OF NORMAL VARIABLES

BU-834-M

by

February, 1984

S. R. Searle
Biometrics Unit, Cornell University, Ithaca, New York

Abstract

Detailed proofs are given of necessary and sufficient conditions for quadratic forms in normal variables (i) to be distributed as χ^2 , (ii) to be independent of each other and (iii) to be independent of a linear form.

1. Introduction

1.1. **Normality**

A vector of variables \underline{x} , of order N , having a normal distribution with mean $\underline{\mu}$ and dispersion matrix \underline{V} , can be represented as

$$\underline{x} \sim n(\underline{\mu}, \underline{V}) ;$$

and a quadratic form of these variables is

$$\underline{x}'\underline{A}\underline{x} \text{ for } \underline{A} \text{ being symmetric, } \underline{A} = \underline{A}' . \quad (1)$$

1.2. Matrix results

\underline{V} in (1), through being a dispersion matrix, is always positive semi-definite (p.s.d.) or positive definite (p.d.). We confine attention to \underline{V} being p.d. and so

$$\underline{V}^{-1} \text{ exists and } \underline{V} = \underline{T}\underline{T}' \text{ for } \underline{T}^{-1} \text{ existing.} \quad (2)$$

In contrast, although \underline{A} of (2) is always symmetric, so that [see Searle (1982), Sec. 7.6e]

$$\underline{A} = \underline{K}\underline{K}' \text{ for } \underline{K} \text{ of full column rank } r_{\underline{A}} \quad (3)$$

for $r_{\underline{A}}$ being the rank of \underline{A} , \underline{A} is not necessarily p.d. or even p.s.d. When \underline{A} is p.s.d. then [see Searle (1982), equation (35), p. 206] we have:

$$\text{for } \underline{A} \text{ being p.s.d., } \underline{A} = \underline{K}\underline{K}' \quad (4)$$

for \underline{K} of full column rank and \underline{K} being real.

In both (3) and (4), $(\underline{K}'\underline{K})^{-1}$ exists and $\underline{K}(\underline{K}'\underline{K})^{-1}$ is a right inverse of \underline{K}' ; but in (4) it is \underline{K} being real that distinguishes (4) from (3), and it is a feature which arises in the sequel.

Use is also made of the results (Searle, 1982, p. 63) that for

$$\underline{A} = \{a_{ij}\}, \quad \text{tr}(\underline{A}\underline{A}') = \sum_{ij} a_{ij}^2$$

and for (5)

$$\underline{A} = \underline{A}' \quad \text{tr}(\underline{A}\underline{A}') = \sum_{ij} a_{ij}^2 = \text{tr}(\underline{A}^2) = \sum_t \lambda_t^2$$

where λ_t is an eigenvalue of \underline{A} .

1.3. Three theorems

The purpose of these notes is to display detailed proofs of three theorems. The first indicates when $\underline{x}'\underline{A}\underline{x}$ has a non-central χ^2 -distribution,

with such a distribution having degrees of freedom f and non-centrality parameter λ being represented by $\chi^2'(f, \lambda)$; the second and third concern independence of $\underline{x}'\underline{A}\underline{x}$ and $\underline{x}'\underline{B}\underline{x}$, and of $\underline{x}'\underline{A}\underline{x}$ and $\underline{B}\underline{x}$.

Theorem 1: $\underline{x}'\underline{A}\underline{x} \sim \chi^2'(f, \lambda)$ if and only if $\underline{A}\underline{V}$ is idempotent; whereupon $f = r_{\underline{A}}$ and $\lambda = \frac{1}{2}\underline{\mu}'\underline{A}\underline{\mu}$.

Theorem 2: $\underline{x}'\underline{A}\underline{x}$ and $\underline{x}'\underline{B}\underline{x}$ are independent if and only if $\underline{A}\underline{V}\underline{B} = \underline{0}$ (or, equivalently, $\underline{B}\underline{V}\underline{A} = \underline{0}$).

Theorem 3: $\underline{x}'\underline{A}\underline{x}$ and $\underline{B}\underline{x}$ are independent if and only if $\underline{B}\underline{V}\underline{A} = \underline{0}$ (or, equivalently, $\underline{A}\underline{V}\underline{B}' = \underline{0}$).

Sufficiency proofs for these theorems are relatively straightforward. They are given here for the sake of completeness. Necessity proofs are not widely available; in some cases they are wrong (e.g., Searle, 1971, pages 57 and 59), sketchy (Graybill, 1976) or are omitted (Arnold, 1981). Although for Theorems 1 and 2 these necessity proofs are lengthy, they are given here in full detail, for the sake of completeness and for availability in teaching.

1.4. Moment generating functions

Under broad regularity conditions the moment generating function (m.g.f.) of a function $f(\underline{x})$, to be denoted by $M_{f(\underline{x})}(t) = E(e^{f(\underline{x})t})$ uniquely determines the probability density function of $f(\underline{x})$. We therefore have occasion to make use of certain properties of m.g.f.'s.

The first is that when two variables x_1 and x_2 are independently distributed, then their joint m.g.f. is the product of each individual m.g.f.

$$M_{(x_1, x_2)}(t_1, t_2) = M_{x_1}(t_1)M_{x_2}(t_2). \quad (6)$$

Second, the moment generating function of a non-central χ^2 density is

$$\text{for } y \sim \chi^2(f, \lambda), \quad M_y(t) = (1 - 2t)^{-\frac{1}{2}f} e^{-\lambda[1-(1-2t)^{-1}]} . \quad (7)$$

Third, for

$$\underline{x} \sim n(\underline{\mu}, \underline{V})$$

$$M_{\underline{x}'\underline{A}\underline{x}}(t) = |\underline{I} - 2t\underline{A}\underline{V}|^{-\frac{1}{2}} \exp\{-\frac{1}{2}\underline{\mu}'[\underline{I} - (\underline{I} - 2t\underline{A}\underline{V})^{-1}]\underline{V}^{-1}\underline{\mu}\} . \quad (8)$$

2. Theorem 1

Sufficiency: that if $\underline{A}\underline{V}$ is idempotent, $\underline{x}'\underline{A}\underline{x} \sim \chi^2(r_{\underline{A}}, \frac{1}{2}\underline{\mu}'\underline{A}\underline{\mu})$.

Proof: Using the non-singularity of \underline{V} , from (3) gives

$$r_{\underline{A}\underline{V}} = r_{\underline{A}} = r, \text{ say.}$$

Therefore, starting with $\underline{A}\underline{V}$ idempotent, we have

$$\underline{A}\underline{V} \text{ has } r \text{ eigenvalues of unity, all others zero.} \quad (9)$$

Hence, with λ_1 being an eigenvalue of $\underline{A}\underline{V}$, the m.g.f. of $\underline{x}'\underline{A}\underline{x}$ in (8) can be written as

$$\begin{aligned} M_{\underline{x}'\underline{A}\underline{x}} &= \prod_{i=1}^n (1 - 2t\lambda_i)^{-\frac{1}{2}} \exp - \frac{1}{2}\underline{\mu}' \left\{ \underline{I} - \left[\underline{I} + \sum_{k=1}^{\infty} (2t)^k (\underline{A}\underline{V})^k \right] \right\} \underline{V}^{-1} \underline{\mu} \\ &= \prod_{i=1}^r (1 - 2t)^{-\frac{1}{2}} \exp - \frac{1}{2}\underline{\mu}' \left[- \sum_{k=1}^{\infty} (2t)^k \underline{A}\underline{V} \right] \underline{V}^{-1} \underline{\mu}, \text{ using (9),} \\ &= (1 - 2t)^{-r/2} \exp - \frac{1}{2}\underline{\mu}' \underline{A}\underline{\mu} \left[- \sum_{k=1}^{\infty} (2t)^k \right] \\ &= (1 - 2t)^{-r/2} \exp - \frac{1}{2}\underline{\mu}' \underline{A}\underline{\mu} [1 - (1 - 2t)^{-1}] . \end{aligned} \quad (10)$$

Comparing (10) with (7) shows that

$$\underline{x}'\underline{A}\underline{x} \sim \chi^2(r, \frac{1}{2}\underline{\mu}'\underline{A}\underline{\mu}) . \quad \text{QED}$$

At least one form of the necessity proof of Theorem 1 needs the preliminaries of the following lemma and corollary.

Lemma: α being algebraic (i.e., a ratio of polynomials) implies that e^α is transcendental.

Proof: Lange (1965).

Corollary: When $P(t)$, $Q(t)$, $R(t)$ and $S(t)$ are polynomials in t , and when $P(t)/Q(t)$ and $R(t)/S(t)$ are rational functions, then

$$P(t)/Q(t) = e^{R(t)/S(t)} \quad \forall t \quad (11)$$

$$\text{implies} \quad R(t) = 0 \quad \text{and} \quad P(t)/Q(t) = 1. \quad (12)$$

Proof: Laha (1956). The crux of the proof is that by the lemma $e^{R(t)/S(t)}$ is transcendental; but $P(t)/Q(t)$ is not. This inconsistency applied to (11) is avoided only by the results in (12).

Necessity part of Theorem 1: that $\underline{x}' \underline{A} \underline{x} \sim \chi^2(f, \lambda)$ implies that $\underline{A} \underline{V}$ is idempotent and $f = r_{\underline{A}}$ and $\lambda = \frac{1}{2} \underline{\mu}' \underline{A} \underline{\mu}$.

Proof: Given that $\underline{x}' \underline{A} \underline{x} \sim \chi^2(f, \lambda)$ we equate (7) and (8) and get, using $\underline{W} = \underline{I} - 2t \underline{A} \underline{V}$,

$$\begin{aligned} \frac{(1 - 2t)^{\frac{1}{2}f}}{|\underline{I} - 2t \underline{A} \underline{V}|^{\frac{1}{2}}} &= \exp \left[\lambda \left(1 - \frac{1}{1 - 2t} \right) - \frac{1}{2} \underline{\mu}' (\underline{I} - \underline{W}^{-1}) \underline{V}^{-1} \underline{\mu} \right] \\ &= \exp \left[\frac{-4\lambda t - (1 - 2t) \underline{\mu}' (\underline{I} - \underline{W}^{-1}) \underline{V}^{-1} \underline{\mu}}{2(1 - 2t)} \right]; \end{aligned}$$

i.e.,

$$\frac{(1 - 2t)^{\frac{1}{2}f}}{|\underline{I} - 2t \underline{A} \underline{V}|^{\frac{1}{2}}} = \exp \left[\frac{-4\lambda t |\underline{W}| - (1 - 2t) \underline{\mu}' (|\underline{W}| \underline{I} - \text{adj} \underline{W}) \underline{V}^{-1} \underline{\mu}}{2(1 - 2t) |\underline{W}|} \right] \quad (13)$$

where $\text{adj} \underline{W}$ is the adjugate matrix of \underline{W} .

Now the left-hand side of (13) is a ratio of two polynomials in t , of order $\frac{1}{2}f$ in the numerator and $\frac{1}{2}n$ in the denominator. And the exponent on the right is also a ratio of two polynomials, of order $n + 1$ in both numerator and denominator. Hence from the preceding corollary, (12) applies and so

$$(1 - 2t)^{\frac{1}{2}f} = |\underline{I} - 2t\underline{AV}|^{\frac{1}{2}}.$$

Write u for $2t$. Then

$$\begin{aligned} (1 - u)^f &= |\underline{I} - u\underline{AV}| \\ &= \prod_{i=1}^n (1 - u\lambda_i) \quad \text{for } \lambda_i \text{ being an eigenvalue of } \underline{AV} \\ &= 1 - u\Sigma\lambda_i + u^2\Sigma\lambda_i\lambda_j - \dots (-1)^k u^k \Sigma\lambda_{i_1}\lambda_{i_2}\dots\lambda_{i_k} + u^n \Pi\lambda_i. \end{aligned} \quad (14)$$

This is an identity in u , with no powers of u greater than u^f on the left-hand side. The same must also be true for the right-hand side. Hence at least one λ_i is zero. Let it be λ_n . Then (14) becomes

$$(1 - u)^f = \prod_{i=1}^{n-1} (1 - u\lambda_i).$$

This argument repeats until we have $n - f$ of the eigenvalues zero and

$$(1 - u)^f = \prod_{i=1}^f (1 - u\lambda_i).$$

From this

$$f \log(1 - u) = \sum_{i=1}^f \log(1 - u\lambda_i)$$

$$f(-u - u^2/2 - u^3/3 - \dots) = -u \sum_{i=1}^f \lambda_i - u^2 \sum_{i=1}^f \lambda_i^2/2 - u^3 \sum_{i=1}^f \lambda_i^3/3 - \dots.$$

Equating powers of u gives

$$f = \sum_{i=1}^f \lambda_i = \sum_{i=1}^f \lambda_i^2 = \dots = \sum_{i=1}^f \lambda_i^f .$$

These equations have a unique solution (see Searle, 1968) which is $\lambda_i = 1$ for $i = 1, 2, \dots, f$. Thus f of the n eigenvalues of $\underline{\underline{AV}}$ are unity, and $n - f$ are zero. Furthermore, $\underline{\underline{A}}$ is symmetric and $\underline{\underline{V}}$ is positive definite and therefore $\underline{\underline{AV}}$ is idempotent (Searle, 1971, p. 37, lemma 9). This being so, (8) reduces to (10) and on comparison with (7) $f = r_{\underline{\underline{A}}}$ and $\lambda = \frac{1}{2} \underline{\underline{\mu}}' \underline{\underline{A}} \underline{\underline{\mu}}$, just as in the sufficiency proof, and so the necessity proof is complete. QED

3. Theorem 2

Sufficiency: that if $\underline{\underline{AVB}} = \underline{\underline{0}}$, then $\underline{\underline{x}}' \underline{\underline{Ax}}$ and $\underline{\underline{x}}' \underline{\underline{Bx}}$ are independent.

Proof: From (3) we have $\underline{\underline{A}} = \underline{\underline{KK}}'$; and similarly $\underline{\underline{B}} = \underline{\underline{LL}}'$ for $\underline{\underline{L}}$ of full column rank. Therefore

$$\underline{\underline{AVB}} = \underline{\underline{0}} \text{ implies } \underline{\underline{KK}}' \underline{\underline{VLL}}' = \underline{\underline{0}} \text{ implies } \underline{\underline{K}}' \underline{\underline{VL}} = \underline{\underline{0}} , \quad (15)$$

this last equality coming from pre- and post-multiplying its predecessor by $(\underline{\underline{K}}' \underline{\underline{K}})^{-1} \underline{\underline{K}}'$ and $\underline{\underline{L}}(\underline{\underline{L}}' \underline{\underline{L}})^{-1}$, respectively. But $\text{cov}(\underline{\underline{K}}' \underline{\underline{x}}, \underline{\underline{x}}' \underline{\underline{L}}) = \underline{\underline{K}}' \underline{\underline{VL}}$. Hence $\underline{\underline{AVB}} = \underline{\underline{0}}$ implies $\text{cov}(\underline{\underline{K}}' \underline{\underline{x}}, \underline{\underline{x}}' \underline{\underline{L}}) = \underline{\underline{0}}$. But $\underline{\underline{K}}' \underline{\underline{x}}$ and $\underline{\underline{x}}' \underline{\underline{L}}$ are vectors of normally distributed random variables, and so their covariance being null implies their independence; i.e., $\underline{\underline{K}}' \underline{\underline{x}}$ and $\underline{\underline{x}}' \underline{\underline{L}}$ are independent. Therefore $\underline{\underline{x}}' \underline{\underline{KK}}' \underline{\underline{x}}$ and $\underline{\underline{x}}' \underline{\underline{LL}}' \underline{\underline{x}}$ are independent; i.e., $\underline{\underline{x}}' \underline{\underline{Ax}}$ and $\underline{\underline{x}}' \underline{\underline{Bx}}$ are independent. QED

Necessity: that $\underline{\underline{x}}' \underline{\underline{Ax}}$ and $\underline{\underline{x}}' \underline{\underline{Bx}}$ being independent implies $\underline{\underline{AVB}} = \underline{\underline{0}}$.

An erroneous attempt at proof

Independence of $\underline{\underline{x}}' \underline{\underline{Ax}}$ and $\underline{\underline{x}}' \underline{\underline{Bx}}$ implies

$$\text{cov}(\underline{\underline{x}}' \underline{\underline{Ax}}, \underline{\underline{x}}' \underline{\underline{Bx}}) = 0 ,$$

i.e.,

$$v(\underline{x}'\underline{A}\underline{x} + \underline{x}'\underline{B}\underline{x}) - v(\underline{x}'\underline{A}\underline{x}) - v(\underline{x}'\underline{B}\underline{x}) = 0 .$$

$$\begin{aligned} \therefore 2\text{tr}[(\underline{A} + \underline{B})\underline{V}]^2 + 4\mu'(\underline{A} + \underline{B})\underline{V}(\underline{A} + \underline{B})\underline{\mu} \\ - 2\text{tr}(\underline{A}\underline{V})^2 - 4\mu'\underline{A}\underline{V}\underline{\mu} - 2\text{tr}(\underline{B}\underline{V})^2 - 4\mu'\underline{B}\underline{V}\underline{\mu} = 0 . \end{aligned}$$

This reduces to

$$\text{tr}(\underline{A}\underline{V}\underline{B}\underline{V}) + 2\mu'\underline{A}\underline{V}\underline{B}\underline{\mu} = 0 . \quad (16)$$

Searle (1981, p. 59) then argues that (16) is to be true for all $\underline{\mu}$ and therefore for $\underline{\mu} = \underline{0}$ and so therefore

$$\text{tr}(\underline{A}\underline{V}\underline{B}\underline{V}) = 0 ; \quad (17)$$

and that (17) then implies $\underline{A}\underline{V}\underline{B} = \underline{0}$. It is this last argument which is fallacious.

A and B being p.s.d.

We use $\underline{V} = \underline{T}\underline{T}'$ of (2) and $\underline{A} = \underline{K}\underline{K}'$ from (3) and similarly $\underline{B} = \underline{L}\underline{L}'$ as in the sufficiency proof. Then, for $\underline{H} = \{h_{ij}\} = \underline{L}'\underline{T}\underline{T}'\underline{K}$

$$\begin{aligned} \text{tr}(\underline{A}\underline{V}\underline{B}\underline{V}) &= \text{tr}(\underline{K}\underline{K}'\underline{T}\underline{T}'\underline{L}\underline{L}'\underline{T}\underline{T}') = \text{tr}[(\underline{L}'\underline{T}\underline{T}'\underline{K})'(\underline{L}'\underline{T}\underline{T}'\underline{K})] \\ &= \text{tr}(\underline{H}'\underline{H}) = \sum_i \sum_j h_{ij}^2 , \end{aligned} \quad (18)$$

using (5). But now, if \underline{A} and \underline{B} are both p.s.d., then \underline{K} and \underline{L} are real, as in (4), and so is every h_{ij} of (18). Hence $\text{tr}(\underline{A}\underline{V}\underline{B}\underline{V}) = 0$ of (17) implies $h_{ij} = 0 \forall i \text{ and } j$, i.e., $\underline{L}'\underline{T}\underline{T}'\underline{K} = \underline{0}$. Therefore $\underline{L}\underline{L}'\underline{T}\underline{T}'\underline{K}\underline{K}' = \underline{0}$, i.e., $\underline{A}\underline{V}\underline{B} = \underline{0}$.

Using cov($\underline{x}'\underline{A}\underline{x}$, $\underline{x}'\underline{B}\underline{x}$)

Independence of two variables always implies their covariance is zero. But the converse holds only for normally distributed variables. This is

well-known but is readily overlooked. It is important here, because the sufficiency proof utilized the fact that $\text{cov}(\underline{\underline{K}}'\underline{\underline{x}}, \underline{\underline{x}}'\underline{\underline{L}}) = 0$ implies $\underline{\underline{K}}'\underline{\underline{x}}$ and $\underline{\underline{x}}'\underline{\underline{L}}$ being independent, since $\underline{\underline{K}}'\underline{\underline{x}}$ and $\underline{\underline{x}}'\underline{\underline{L}}$ are normally distributed.

For the necessity proof we start with $\underline{\underline{x}}'\underline{\underline{A}}\underline{\underline{x}}$ and $\underline{\underline{x}}'\underline{\underline{B}}\underline{\underline{x}}$ being independent. Therefore $\text{cov}(\underline{\underline{x}}'\underline{\underline{A}}\underline{\underline{x}}, \underline{\underline{x}}'\underline{\underline{B}}\underline{\underline{x}}) = 0$. But, although, as we will show,

$$\underline{\underline{x}}'\underline{\underline{A}}\underline{\underline{x}} \text{ and } \underline{\underline{x}}'\underline{\underline{B}}\underline{\underline{x}} \text{ independent implies } \underline{\underline{A}}\underline{\underline{V}}\underline{\underline{B}} = 0$$

and

$$\underline{\underline{x}}'\underline{\underline{A}}\underline{\underline{x}} \text{ and } \underline{\underline{x}}'\underline{\underline{B}}\underline{\underline{x}} \text{ independent implies } \text{cov}(\underline{\underline{x}}'\underline{\underline{A}}\underline{\underline{x}}, \underline{\underline{x}}'\underline{\underline{B}}\underline{\underline{x}}) = 0 ,$$

we can illustrate that, in general

$$\text{cov}(\underline{\underline{x}}'\underline{\underline{A}}\underline{\underline{x}}, \underline{\underline{x}}'\underline{\underline{B}}\underline{\underline{x}}) = 0 \text{ does not imply } \underline{\underline{A}}\underline{\underline{V}}\underline{\underline{B}} = 0 . \quad (19)$$

Example of (19)

$$\text{Suppose } \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \sim \eta \left[\begin{bmatrix} 0 \\ 0 \end{bmatrix} , \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right] .$$

Let

$$x_1^2 + x_2^2 = \underline{\underline{x}}'\underline{\underline{A}}\underline{\underline{x}} \quad \text{for } \underline{\underline{A}} = \underline{\underline{I}}$$

and

$$x_1^2 - x_2^2 = \underline{\underline{x}}'\underline{\underline{B}}\underline{\underline{x}} \quad \text{for } \underline{\underline{B}} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} .$$

Then

$$\text{cov}(x_1^2 + x_2^2, x_1^2 - x_2^2) = v(x_1^2) - v(x_2^2) = v(x_1^2) - v(x_1^2) = 0 .$$

But

$$\underline{\underline{A}}\underline{\underline{V}}\underline{\underline{B}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \neq 0 ;$$

i.e., $\text{cov}(\underline{x}'\underline{A}\underline{x}, \underline{x}'\underline{B}\underline{x}) = 0$ but $\underline{A}\underline{V}\underline{B} \neq 0$.

We are driven to the conclusion that the necessity proof demands using m.g.f.'s.

Moment generating functions

Given the independence of $\underline{x}'\underline{A}\underline{x}$ and $\underline{x}'\underline{B}\underline{x}$ we use (6) to write

$$M_{\underline{x}'\underline{A}\underline{x}, \underline{x}'\underline{B}\underline{x}}(t, u) = M_{\underline{x}'\underline{A}\underline{x}}(t)M_{\underline{x}'\underline{B}\underline{x}}(u) .$$

Then using (8) for each M gives

$$\begin{aligned} & \frac{\exp - \frac{1}{2}\underline{\mu}'[\underline{I} - (\underline{I} - 2t\underline{A}\underline{V} - 2u\underline{B}\underline{V})^{-1}]\underline{V}^{-1}\underline{\mu}}{|\underline{I} - 2t\underline{A}\underline{V} - 2u\underline{B}\underline{V}|^{\frac{1}{2}}} \\ &= \frac{\exp - \frac{1}{2}\underline{\mu}'[\underline{I} - (\underline{I} - 2t\underline{A}\underline{V})^{-1} + \underline{I} - (\underline{I} - 2u\underline{B}\underline{V})^{-1}]\underline{V}^{-1}\underline{\mu}}{|\underline{I} - 2t\underline{A}\underline{V}|^{\frac{1}{2}}|\underline{I} - 2u\underline{B}\underline{V}|^{\frac{1}{2}}} . \end{aligned}$$

Hence

$$\begin{aligned} & \left[\frac{|\underline{I} - 2t\underline{A}\underline{V}||\underline{I} - 2u\underline{B}\underline{V}|}{|\underline{I} - 2t\underline{A}\underline{V} - 2u\underline{B}\underline{V}|} \right]^{\frac{1}{2}} \\ &= \frac{\exp - \frac{1}{2}\underline{\mu}'[\underline{I} - (\underline{I} - 2t\underline{A}\underline{V})^{-1} + \underline{I} - (\underline{I} - 2u\underline{B}\underline{V})^{-1}]\underline{V}^{-1}\underline{\mu}}{\exp - \frac{1}{2}\underline{\mu}'[\underline{I} - (\underline{I} - 2t\underline{A}\underline{V} - 2u\underline{B}\underline{V})^{-1}]\underline{V}^{-1}\underline{\mu}} . \end{aligned}$$

Applying the corollary used in proving the necessity part of Theorem 1, in particular (12) thereof, gives

$$|\underline{I} - 2t\underline{A}\underline{V}||\underline{I} - 2u\underline{B}\underline{V}| = |\underline{I} - 2t\underline{A}\underline{V} - 2u\underline{B}\underline{V}| . \quad (20)$$

The problem now is to show that (20) implies $\underline{A}\underline{V}\underline{B} = 0$. We give three different proofs.

Scarowsky's proof

We can begin the proof in Scarowsky (1973) with (2). Then

$$\begin{aligned}
 |\underline{I} - 2t\underline{A}\underline{V}| &= |\underline{I} - 2t\underline{A}\underline{T}\underline{T}'| = |\underline{T}'^{-1}| |\underline{I} - 2t\underline{T}'\underline{A}\underline{T}| |\underline{T}'| \\
 &= |\underline{I} - 2t\underline{T}'\underline{A}\underline{T}| \\
 &= |\underline{I} - 2t\underline{\dot{A}}| \quad (21)
 \end{aligned}$$

for

$$\underline{\dot{A}} = \underline{T}'\underline{A}\underline{T} = \underline{\dot{A}}'. \quad \text{Similarly for } \underline{\dot{B}} = \underline{T}'\underline{B}\underline{T} = \underline{\dot{B}}' \quad (22)$$

and then (20) is

$$|\underline{I} - 2t\underline{\dot{A}}| |\underline{I} - 2u\underline{\dot{B}}| = |\underline{I} - 2t\underline{\dot{A}} - 2u\underline{\dot{B}}|. \quad (23)$$

For any matrix \underline{P} having real eigenvalues λ_i (as do $\underline{\dot{A}}$ and $\underline{\dot{B}}$ because they are symmetric)

$$\begin{aligned}
 \log |\underline{I} - \underline{P}| &= \log \prod_i (1 - \lambda_i) = \sum_i \log(1 - \lambda_i) \\
 &= -\sum_i \sum_{k=1}^{\infty} \lambda_i^k / k = -\sum_{k=1}^{\infty} (\sum_i \lambda_i^k) / k \\
 &= -\sum_{k=1}^{\infty} \text{tr}(\underline{P}^k) / k.
 \end{aligned}$$

Using this in (23) gives

$$\sum_{k=1}^{\infty} t^k \text{tr}(\underline{\dot{A}}^k) / k + \sum_{k=1}^{\infty} u^k \text{tr}(\underline{\dot{B}}^k) / k = \sum_{k=1}^{\infty} \text{tr}(t\underline{\dot{A}} + u\underline{\dot{B}})^k / k.$$

Equate coefficients of $t^2 u^2$:

$$\begin{aligned}
 0 &= \text{coefficient of } t^2 u^2 \text{ in } \frac{1}{4} \text{tr}(t\underline{\dot{A}} + u\underline{\dot{B}})^4 \\
 &= \frac{1}{4} \text{tr}(\underline{\dot{A}}\underline{\dot{A}}\underline{\dot{B}}\underline{\dot{B}} + \underline{\dot{A}}\underline{\dot{B}}\underline{\dot{A}}\underline{\dot{B}} + \underline{\dot{B}}\underline{\dot{A}}\underline{\dot{A}}\underline{\dot{B}} + \underline{\dot{B}}\underline{\dot{B}}\underline{\dot{A}}\underline{\dot{A}} + \underline{\dot{B}}\underline{\dot{A}}\underline{\dot{B}}\underline{\dot{A}} + \underline{\dot{A}}\underline{\dot{B}}\underline{\dot{B}}\underline{\dot{A}}) \\
 &= \frac{1}{4} [4 \text{tr}(\underline{\dot{A}}^2 \underline{\dot{B}}^2) + 2 \text{tr}(\underline{\dot{A}}\underline{\dot{B}})^2].
 \end{aligned}$$

Therefore, because $\dot{\underline{\underline{A}}}$ and $\dot{\underline{\underline{B}}}$ are symmetric, this is

$$0 = \text{tr}(\dot{\underline{\underline{A}}}\dot{\underline{\underline{B}}} + \dot{\underline{\underline{B}}}\dot{\underline{\underline{A}}})(\dot{\underline{\underline{A}}}\dot{\underline{\underline{B}}} + \dot{\underline{\underline{B}}}\dot{\underline{\underline{A}}})' + 2\text{tr}\dot{\underline{\underline{A}}}\dot{\underline{\underline{B}}}(\dot{\underline{\underline{A}}}\dot{\underline{\underline{B}}})' . \quad (24)$$

$\underline{\underline{V}}$ is p.d., $\underline{\underline{T}}$ is real, and so every matrix in (24) is real. Furthermore, each term in (24) has the form (5), and so (24) implies

$$\dot{\underline{\underline{A}}}\dot{\underline{\underline{B}}} = 0 . \quad (25)$$

Therefore, from (22), $\underline{\underline{T}}'\underline{\underline{A}}\underline{\underline{T}}\underline{\underline{T}}'\underline{\underline{B}}\underline{\underline{T}} = 0$; and since $\underline{\underline{T}}$ is non-singular, and $\underline{\underline{V}} = \underline{\underline{T}}\underline{\underline{T}}'$, this means $\underline{\underline{A}}\underline{\underline{V}}\underline{\underline{B}} = 0$. QED

Guttman's proof

This proof is Guttman's (1982, p. 83) version of a proof by Lancaster (1954) of the theorem as stated by Craig (1943).

In (23), rewrite $2t$ and $2u$ as t and u , respectively, so that (23) becomes

$$|\underline{\underline{I}} - t\underline{\underline{A}}||\underline{\underline{I}} - u\underline{\underline{B}}| = |\underline{\underline{I}} - t\underline{\underline{A}} - u\underline{\underline{B}}| . \quad (26)$$

For orthogonal $\underline{\underline{U}}$ define $\underline{\underline{U}}$ such that

$$\underline{\underline{U}}'\dot{\underline{\underline{A}}}\underline{\underline{U}} = \underline{\underline{C}} = \begin{bmatrix} \underline{\underline{D}} & \underline{\underline{O}} \\ \underline{\underline{O}} & \underline{\underline{O}} \end{bmatrix} \quad \text{for} \quad \underline{\underline{D}} = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_r \end{bmatrix} \quad (27)$$

where $\lambda_1, \dots, \lambda_r$ are the $r = r_{\underline{\underline{A}}}$ non-zero eigenvalues of the symmetric matrix $\dot{\underline{\underline{A}}}$. Let

$$\underline{\underline{U}}'\dot{\underline{\underline{B}}}\underline{\underline{U}} = \underline{\underline{G}} = \begin{bmatrix} \underline{\underline{G}}_{11} & \underline{\underline{G}}_{12} \\ \underline{\underline{G}}_{21} & \underline{\underline{G}}_{22} \end{bmatrix} \quad \text{for} \quad \underline{\underline{G}}_{21} = (\underline{\underline{G}}_{12})' . \quad (28)$$

Then, because $|\underline{U}'\underline{U}| = |\underline{U}'||\underline{U}| = 1$, since \underline{U} is orthogonal, (26) gives

$$|\underline{I} - t\underline{C}||\underline{I} - u\underline{G}| = |\underline{I} - t\underline{C} - u\underline{G}| . \quad (29)$$

But

$$\begin{aligned} |\underline{I} - t\underline{C} - u\underline{G}| &= \begin{vmatrix} \underline{I} - t\underline{D} - u\underline{G}_{11} & -u\underline{G}_{12} \\ -u\underline{G}_{21} & \underline{I} - u\underline{G}_{22} \end{vmatrix} \\ &= |\underline{I} - u\underline{G}_{22}||\underline{I} - t\underline{D} - u\underline{G}_{11} - u^2\underline{G}_{12}(\underline{I} - u\underline{G}_{22})^{-1}\underline{G}_{21}| , \end{aligned} \quad (30)$$

using the standard result (Searle, 1982, p. 258) that

$$\begin{vmatrix} \underline{A}_1 & \underline{A}_2 \\ \underline{A}_3 & \underline{A}_4 \end{vmatrix} = |\underline{A}_4| |\underline{A}_1 - \underline{A}_2\underline{A}_4^{-1}\underline{A}_3| .$$

Now equate coefficients of t^r in (30), using the form \underline{C} and \underline{D} in (27):

$$(-1)^r \prod_{i=1}^r \lambda_i |\underline{I} - u\underline{G}| = |\underline{I} - u\underline{G}_{22}| (-1)^r \prod_{i=1}^r \lambda_i .$$

Hence

$$|\underline{I} - u\underline{G}| = |\underline{I} - u\underline{G}_{22}| .$$

Therefore

$$\text{eigenvalues of } \underline{G} = \text{eigenvalues of } \underline{G}_{22} . \quad (31)$$

Applying (5) to (31) gives

$$\sum_{i=1}^n \sum_{j=1}^n g_{ij}^2 = \sum_{i=1}^{n-r} \sum_{j=1}^{n-r} g_{22,ij}^2 , \quad (32)$$

where g_{ij} is a typical element of \underline{G} of order n , and $g_{22,ij}$ is a typical element of \underline{G}_{22} of order $n - r$. But, as in (28), \underline{G}_{22} is a submatrix of \underline{G} . Hence (32) implies

$$\sum_{i=1}^r \sum_{j=1}^r g_{11,ij}^2 = 0 \quad \text{and} \quad \sum_{i=1}^r \sum_{j=1}^{n-r} g_{12,ij}^2 = 0$$

and so

$$\underline{G}_{11} = \underline{0} \quad \text{and} \quad \underline{G}_{12} = \underline{0} . \quad (33)$$

Hence

$$\underline{G} = \begin{bmatrix} 0 & 0 \\ 0 & \underline{G}_{22} \end{bmatrix} . \quad (34)$$

Therefore

$$\underline{CG} = \begin{bmatrix} \underline{D} & \underline{0} \\ \underline{0} & \underline{0} \end{bmatrix} \begin{bmatrix} \underline{0} & \underline{0} \\ \underline{0} & \underline{G}_{22} \end{bmatrix} = \underline{0} ,$$

i.e.,

$$\underline{U}' \underline{\dot{A}} \underline{U} \underline{U}' \underline{\dot{B}} \underline{U} = \underline{0} ,$$

$$\underline{\dot{A}} \underline{\dot{B}} = \underline{0}$$

and so, as following (25), $\underline{A} \underline{V} \underline{B} = \underline{0}$. QED

Krafft's proof

Using (16) and (17), Krafft (1978) starts with

$$\underline{\mu}' \underline{A} \underline{V} \underline{B} \underline{\mu} = 0 \quad \forall \underline{\mu} . \quad (35)$$

Writing $\underline{\mu}$ as $\underline{p} + \underline{q}$ this is

$$(\underline{p}' + \underline{q}') \underline{A} \underline{V} \underline{B} (\underline{p} + \underline{q}) = 0 \quad \forall \underline{p} \text{ and } \underline{q} .$$

Thus

$$\underline{p}' \underline{A} \underline{V} \underline{B} \underline{p} + \underline{q}' \underline{A} \underline{V} \underline{B} \underline{q} + \underline{p}' \underline{A} \underline{V} \underline{B} \underline{q} + \underline{q}' \underline{A} \underline{V} \underline{B} \underline{p} = 0 . \quad (36)$$

But on applying (35) to the first two terms (36) becomes, on using $(\underline{\underline{AVB}})' = \underline{\underline{BVA}}$,

$$\underline{\underline{p}}'(\underline{\underline{AVB}} + \underline{\underline{BVA}})\underline{\underline{q}} = 0 \quad \forall \underline{\underline{p}} \text{ and } \underline{\underline{q}} . \quad (37)$$

Letting $\underline{\underline{p}}$ and $\underline{\underline{q}}$ be, in turn, the columns of an identity matrix, (37) yields

$$\underline{\underline{AVB}} + \underline{\underline{BVA}} = \underline{\underline{0}} ;$$

i.e.,

$$\underline{\underline{AVB}} = -\underline{\underline{BVA}} . \quad (38)$$

Now in (23) let $\underline{\underline{u}} = \underline{\underline{t}}$. Then

$$|\underline{\underline{I}} - \underline{\underline{t}}\underline{\underline{\dot{A}}}| |\underline{\underline{I}} - \underline{\underline{t}}\underline{\underline{\dot{B}}}| = |\underline{\underline{I}} - \underline{\underline{t}}(\underline{\underline{\dot{A}}} + \underline{\underline{\dot{B}}})| . \quad (39)$$

On the left-hand side of (39) each determinant is a polynomial in $\underline{\underline{t}}$, of order $r_{\underline{\underline{A}}}$ for one, and $r_{\underline{\underline{B}}}$ for the other; and the right-hand side is a polynomial of order $r_{\underline{\underline{A+B}}}$. This is so because $\underline{\underline{\dot{A}}}$ and $\underline{\underline{\dot{B}}}$ are symmetric and their number of non-zero eigenvalues is their ranks. Thus

$$r_{\underline{\underline{A}}} + r_{\underline{\underline{B}}} = r_{\underline{\underline{A+B}}} . \quad (40)$$

One product of (27) and (28) is

$$\underline{\underline{U}}'\underline{\underline{\dot{A}}}\underline{\underline{U}}\underline{\underline{U}}'\underline{\underline{\dot{B}}}\underline{\underline{U}} = \begin{bmatrix} \underline{\underline{D}} & \underline{\underline{0}} \\ \underline{\underline{0}} & \underline{\underline{0}} \end{bmatrix} \begin{bmatrix} \underline{\underline{G}}_{11} & \underline{\underline{G}}_{12} \\ \underline{\underline{G}}_{21} & \underline{\underline{G}}_{22} \end{bmatrix} .$$

Using the orthogonality of $\underline{\underline{U}}$, equation (22) for $\underline{\underline{\dot{A}}}$ and $\underline{\underline{\dot{B}}}$, and $\underline{\underline{TT}}' = \underline{\underline{V}}$ from (2), this is

$$\underline{\underline{U}}'\underline{\underline{T}}'\underline{\underline{AVBTU}} = \begin{bmatrix} \underline{\underline{DG}}_{11} & \underline{\underline{DG}}_{12} \\ \underline{\underline{0}} & \underline{\underline{0}} \end{bmatrix} . \quad (41)$$

Similarly, the other product of (27) and (28) is

$$\underline{\underline{U}}' \underline{\underline{B}} \underline{\underline{U}} \underline{\underline{A}} \underline{\underline{U}} = \begin{bmatrix} \underline{\underline{G}}_{11} & \underline{\underline{G}}_{12} \\ \underline{\underline{G}}_{21} & \underline{\underline{G}}_{22} \end{bmatrix} \begin{bmatrix} \underline{\underline{D}} & \underline{\underline{0}} \\ \underline{\underline{0}} & \underline{\underline{0}} \end{bmatrix}$$

so that

$$-\underline{\underline{U}}' \underline{\underline{T}}' \underline{\underline{B}} \underline{\underline{V}} \underline{\underline{A}} \underline{\underline{U}} = - \begin{bmatrix} \underline{\underline{G}}_{11} \underline{\underline{D}} & \underline{\underline{0}} \\ \underline{\underline{G}}_{21} \underline{\underline{D}} & \underline{\underline{0}} \end{bmatrix} \quad (42)$$

But (38) indicates that (41) and (42) are equal:

$$\begin{bmatrix} \underline{\underline{D}} \underline{\underline{G}}_{11} & \underline{\underline{D}} \underline{\underline{G}}_{12} \\ \underline{\underline{0}} & \underline{\underline{0}} \end{bmatrix} = - \begin{bmatrix} \underline{\underline{G}}_{11} \underline{\underline{D}} & \underline{\underline{0}} \\ \underline{\underline{G}}_{21} \underline{\underline{D}} & \underline{\underline{0}} \end{bmatrix}.$$

Therefore $\underline{\underline{G}}_{12} = \underline{\underline{0}}$; and so, in (28),

$$\underline{\underline{U}}' \underline{\underline{B}} \underline{\underline{U}} = \underline{\underline{G}} = \begin{bmatrix} \underline{\underline{G}}_{11} & \underline{\underline{0}} \\ \underline{\underline{0}} & \underline{\underline{G}}_{22} \end{bmatrix} \quad (43)$$

and in (41)

$$\underline{\underline{U}}' \underline{\underline{T}}' \underline{\underline{A}} \underline{\underline{V}} \underline{\underline{B}} \underline{\underline{U}} = \begin{bmatrix} \underline{\underline{D}} \underline{\underline{G}}_{11} & \underline{\underline{0}} \\ \underline{\underline{0}} & \underline{\underline{0}} \end{bmatrix}.$$

Therefore from (27) and (43)

$$\underline{\underline{U}}' \underline{\underline{A}} \underline{\underline{U}} + \underline{\underline{U}}' \underline{\underline{B}} \underline{\underline{U}} = \begin{bmatrix} \underline{\underline{D}} + \underline{\underline{G}}_{11} & \underline{\underline{0}} \\ \underline{\underline{0}} & \underline{\underline{G}}_{22} \end{bmatrix}. \quad (44)$$

Now from (44)

$$r_{\underline{A+B}} = r_{\underline{D+G_{11}}} + r_{\underline{G_{22}}} ;$$

and on using (40) this is

$$r_{\underline{A}} + r_{\underline{B}} = r_{\underline{D+G_{11}}} + r_{\underline{G_{22}}} . \quad (45)$$

But (27) implies $r_{\underline{A}} = r_{\underline{D}}$ and (43) gives $r_{\underline{B}} = r_{\underline{G_{11}}} + r_{\underline{G_{22}}}$. Therefore (45) is

$$r_{\underline{D}} + r_{\underline{G_{11}}} + r_{\underline{G_{22}}} = r_{\underline{D+G_{11}}} + r_{\underline{G_{22}}} ,$$

i.e.,

$$r_{\underline{D}} + r_{\underline{G_{11}}} = r_{\underline{D+G_{11}}} . \quad (46)$$

But \underline{D} is diagonal of full rank. Therefore

$$r_{\underline{D+G_{11}}} \leq r_{\underline{D}} ,$$

so that from (46)

$$r_{\underline{D}} + r_{\underline{G_{11}}} \leq r_{\underline{D}} .$$

Hence

$$r_{\underline{G_{11}}} = 0 ,$$

i.e.,

$$\underline{G_{11}} = 0 .$$

Thus in (43)

$$\underline{G} = \begin{bmatrix} \underline{0} & \underline{0} \\ \underline{0} & \underline{G_{22}} \end{bmatrix}$$

and the rest of the proof follows as from (34). QED

4. Theorem 3

Sufficiency: that $\underline{\underline{BVA}} = \underline{\underline{0}}$ implies $\underline{\underline{x'Ax}}$ and $\underline{\underline{Bx}}$ are independent.

Proof: Using $\underline{\underline{A}} = \underline{\underline{KK'}}$ of (3), we have

$$\underline{\underline{BVA}} = \underline{\underline{0}} \text{ implies } \underline{\underline{BVKK'}} = \underline{\underline{0}} \text{ implies } \underline{\underline{BVK}} = \underline{\underline{0}} ,$$

this last equality coming from post-multiplying $\underline{\underline{BVKK'}}$ by $\underline{\underline{K(K'K)^{-1}}}$. But $\text{cov}(\underline{\underline{Bx}}, \underline{\underline{x'K}}) = \underline{\underline{BVK}}$. Hence $\underline{\underline{BVA}} = \underline{\underline{0}}$ implies $\text{cov}(\underline{\underline{Bx}}, \underline{\underline{x'K}}) = \underline{\underline{0}}$. But $\underline{\underline{Bx}}$ and $\underline{\underline{x'K}}$ are vectors of normally distributed random variables and so their covariance being null implies their independence; i.e., $\underline{\underline{Bx}}$ and $\underline{\underline{x'K}}$ are independent. Therefore $\underline{\underline{Bx}}$ and $\underline{\underline{x'KK'x}} = \underline{\underline{x'Ax}}$ are independent. QED

Necessity: that $\underline{\underline{x'Ax}}$ and $\underline{\underline{Bx}}$ being independent implies $\underline{\underline{BVA}} = \underline{\underline{0}}$.

Proof (adapted from Krafft, 1978): When $\underline{\underline{Bx}}$ is independent of $\underline{\underline{x'Ax}}$, then so is $\underline{\underline{x'B'Bx}}$. Therefore, by Theorem 2

$$\underline{\underline{AVB'B}} = \underline{\underline{0}} .$$

Hence, using (7) on page 63 of Searle (1982)

$$\underline{\underline{AVB'}} = \underline{\underline{0}} ,$$

i.e.,

$$\underline{\underline{BVA}} = \underline{\underline{0}} . \quad \text{QED}$$

5. Singular V

Adaptations and generalizations of Theorems 1, 2 and 3 to the case when $\underline{\underline{V}}$ is singular are available in the literature. Nagase and Banerjee (1976) and Searle (1971) are at least two places that will provide the interested reader with entr   to the relevant literature.

References

- Arnold, S. F. (1981). *The Theory of Linear Models and Multivariate Analysis*. Wiley, New York.
- Craig, A. T. (1943). Note on the independence of certain quadratic forms. *Ann. Math. Statist.* 14, 195-197.
- Graybill, F. A. (1976). *Theory and Applications of the Linear Model*. Duxbury, N. Scituate, Mass.
- Guttman, Irwin (1982). *Linear Models: An Introduction*. Wiley, New York.
- Krafft, Olaf (1978). *Lineare Statistische Modelle und Optimale Versuchspläne*. Vandenhoeck and Ruprecht in Göttingen und Zurich.
- Laha, R. G. (1956). On the stochastic independence of two second-degree polynomial statistics in normally distributed variables. *Ann. Math. Statist.* 27, 790-796.
- Lancaster, H. O. (1954). Traces and cumulants of quadratic forms in normal variables. *J. Roy. Statist. Soc. (B)* 16, 247-254.
- Lange, Serge (1965). *Algebra*. Addison-Wesley.
- Nagase, G. and Banerjee, K. S. (1976). Independence of second-degree polynomial statistics in normal variate. *The American Statistician* 30, 201.
- Scarowsky, E. (1973). Quadratic forms in normal variables. Tech. Report, Department of Mathematics, McGill University.
- Searle, S. R. (1968). A remark on solving equations in sums of powers. *J. Roy. Statist. Soc. (B)* 30, 567-569.
- Searle, S. R. (1971). *Linear Models*. Wiley, New York.
- Searle, S. R. (1982). *Matrix Algebra Useful for Statistics*. Wiley, New York.